2019

A trio of sigma factors control hormogonium development in Nostoc punctiforme

Alfonso Gonzalez Jr.
University of the Pacific, a_gonzalez13@u.pacific.edu

A TRIO OF SIGMA FACTORS CONTROL HORMOGONIUM DEVELOPMENT IN *NOSTOC PUNCTIFORME*

by

Alfonso González Diarte

A Thesis Submitted to the

Graduate School

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

College of the Pacific
Biological Sciences

University of the Pacific
Stockton, California

2019

A TRIO OF SIGMA FACTORS CONTROL HORMOGONIUM DEVELOPMENT IN *NOSTOC PUNCTIFORME*

by

Alfonso Gonzalez Diarte

APPROVED BY:

Thesis Advisor: Douglas Risser, Ph.D.

Committee Member: Geoffrey Lin-Cereghino, Ph.D.

Committee Member: Lisa Wrischnik, Ph.D.

Department Chairs: Eric Thomas, Ph.D. & Lisa Wrischnik, Ph.D.

Dean of Graduate School: Thomas Naehr, Ph.D.

DEDICATION

For my parents, Alfonso Gonzalez Razo, and Carmen Diarte Gonzalez, whom have always supported my ambitions and dreams

A Trio of Sigma Factors Control Hormogonium Development in *Nostoc punctiforme*

Abstract

by Alfonso González Diarte

University of the Pacific
2019

Cyanobacteria are prokaryotes capable of oxygenic photosynthesis, and for many species, nitrogen fixation, giving cyanobacteria an important role in global carbon and nitrogen cycles. Furthermore, multicellular filamentous cyanobacteria are developmentally complex, capable of differentiation into different cell types, including cells capable of nitrogen fixation and cells for motility, making them an ideal platform for studying development, as well as for practical use in biotechnology. Understanding how developmental programmes are activated require an understanding of the role of alternative sigma factors, which are required for transcriptional activation in bacteria. In order to investigate the gene regulatory network and to determine the role of alternative sigma factors in hormogonium development, real time PCR and Next Generation RNA-seq were used to measure expression levels of genes involved in hormogonium development and to further characterise the nature of the hormogonium developmental programme in the filamentous cyanobacterium *Nostoc punctiforme*. The results support a model where a hierarchal sigma factor cascade activates hormogonium development, in which expression of *sigJ* activates expression of the sigma factors *sigC* and *sigF*, as well as a wide

range of other genes, including those involved in the type IV pilus (T4P), chemotaxis-like systems, and cell architecture. SigC and SigF have more limited roles: cell division genes are dependent on SigC and *pilA* expression was stringently SigF-dependent. Interestingly, SigC was also found to enhance expression of *sigJ* during hormogonium development, implying a potential positive feedback loop between *sigJ* and *sigC*.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

Chapter 1: Introduction

Background on Bacteria

Bacteria, as most people know, are invisible, microscopic organisms that are found almost everywhere on Earth. Commonly associated with disease, bacteria are one of the most studied organisms due to their impressive diversity. These studies are not only for combating infectious disease, but also for applications in food microbiology such as cheese production, recombinant gene expression, such as the production of human insulin or other drugs, and technologies like CRISPR/Cas9[1]. Because bacteria are extremely diverse, they also play an important role in the environment, including decomposition of organic compounds, whether the bacteria are out in the environment or inside our bodies. They also play a critical role in the shaping of our atmosphere and in the recycling of nutrients in the soil.

Bacteria, like many organisms on Earth, rely on a carbon source, such as glucose, for viability. These carbon sources are found almost anywhere, especially in decaying animal and plant matter. This carbon source is usually fixed, such as in the form of glucose, which is broken down for energy via cellular respiration or fermentation. Non-fixed carbon sources, such as carbon dioxide, are used to produce sugars in photosynthetic organisms, such as plants. Likewise, a group of bacteria are also capable of carrying out such a process, and thus, are capable of producing their own fixed carbon source. In fact, chloroplasts in modern-day plant species are thought to have arisen from these primitive microbes, which were capable of carrying out photosynthesis[2].

These Gram-negative bacteria are known as cyanobacteria, or blue-green algae. They are an extremely old group of microbes, first appearing in the fossil record 3.5 billion years ago. Their ability to perform oxygenic photosynthesis eliminates their reliance on an external source of fixed carbon. Oxygen is produced as a by-product of photosynthesis, and organisms eventually evolved to use oxygen in aerobic respiration. Cyanobacteria are thought to be responsible for the oxygenation of the early Earth's atmosphere around 2.4 billion years ago and in the maintenance of oxygen levels in today's atmosphere[3]. Because of these qualities, cyanobacteria remain a critical part of our ecosystem, as they have throughout much of Earth's history.

Background on Cyanobacteria, Development, and Applications

Cyanobacteria are morphologically diverse, including single-celled cyanobacteria and multi-cellular filamentous strains. Single-celled cyanobacteria are relatively simple while filamentous ones are more developmentally complex. Many of these cyanobacteria are not only capable of photosynthesis but are also capable of fixing atmospheric nitrogen into ammonia, which functions as a precursor to nucleic acids and amino acids. Nitrogen fixation is a unique trait rarely found in other organisms. Cyanobacteria are also capable of differentiating into specialised cells for survival in extreme environments, analogous to bacterial endospores[4], and non-dividing cells that function in motility. Due to these unique characteristics, cyanobacteria have the potential for application in biotechnology.

Differentiation of filamentous cyanobacteria. Multi-cellular filamentous cyanobacteria are very versatile, in part due to the ability of vegetative cells to differentiate into three different cell types in response to changing environmental conditions or signals. Filamentous cyanobacteria generally exist as vegetative cells, which perform photosynthesis and other

ordinary metabolic processes mainly for cell growth and division. Vegetative cells contribute to the viability of the filament and growth of the colony, which both occur when environmental conditions are ideal for growth. This includes a good source of sunlight and carbon dioxide for photosynthesis, and a source of fixed nitrogen, such as ammonium or nitrate for building nitrogen-based molecules such as nucleic and amino acids. The three alternative fates of filamentous cyanobacteria cell differentiation are 1.) heterocysts, which fix nitrogen, 2.) akinetes, whose development are triggered by environmental stress, and 3.) hormogonia, which function in dispersal, and are largely conserved throughout the taxa. However, not all cyanobacteria can differentiate into all of the three cell development fates.

Heterocysts. There are times when fixed nitrogen in the environment may be scarce, and some vegetative cells differentiate into heterocysts. Heterocysts are terminally differentiated cells which function exclusively for nitrogen fixation[5,6]. Photosynthesis does not occur in heterocysts as the enzyme responsible for nitrogen fixation is sensitive to oxygen[7]. Heterocysts instead house the enzyme nitrogenase, which converts the extremely inert dinitrogen into ammonium, which is then used by the rest of the filament for nucleic acid synthesis, protein synthesis, and other nitrogen-dependent metabolic processes. These heterocysts are regularly spaced out, usually at an interval of 10-15 vegetative cells per heterocyst cell, which ensures equal spreading of nitrogen fixation throughout the filament. Heterocyst differentiation is the most studied developmental event in cyanobacteria.

Akinetes. If environmental conditions become more limiting, in which the viability of both vegetative and heterocyst cells is compromised, the cyanobacteria may differentiate instead into akinetes. Akinetes are hardier than vegetative cells and become specialised for survival under more stressful conditions, such as extreme cold[4,8]. Akinetes are generally larger than

vegetative cells and are likely to be seen in cultures during the stationary growth phase. Akinetes can be compared to Gram-positive bacterial endospores, but are not as metabolically quiet, and are not as hardy as endospores. When environmental conditions become more favourable for the microbe, akinetes re-germinate into vegetative cells and resume normal metabolism. Not all cyanobacteria can differentiate into akinetes. In those that do, akinete differentiation has been discovered exclusively in heterocyst-forming bacteria. However, not all heterocyst-forming bacteria can differentiate into akinetes.

Hormogonia. Finally, filamentous cyanobacteria may differentiate into hormogonia, which are motile, non-dividing filaments that allow them to migrate to new habitats and are a prerequisite for establishing symbiosis with new plant partners. As regulation of hormogonium development is a main focus of the study, more information about hormogonium development, morphology, and physiology will be explained below.

Potential applications of cyanobacteria. Bacteria have become useful for biosynthesis of products of value, including proteins, pharmaceutical drugs, foods, and industrial chemicals. Cyanobacteria can also potentially be used for those purposes because cyanobacteria have relatively simple nutritive requirements, are cheap to grow, and can be easily genetically modified like most heterotrophic bacteria[9,10]. Through genetic engineering, photosynthesis may be coupled with recombinant pathways to produce a recombinant product of value in a cost-effective manner. Hormogonium filaments in particular are an attractive platform for industrial use as hormogonia generally allocate their energy into motility rather than biomass production. Therefore, cyanobacteria could potentially be genetically modified so that energy is allocated away from motility and instead used to create a useful recombinant product, which may be used as a biofuel. Additionally, because heterocysts provide a source of fixed nitrogen, and because

cyanobacteria are capable of forming plant-cyanobacteria partnerships, cyanobacteria could potentially be genetically modified to help with crop production, limiting the need for nitrogen-based fertiliser.

Alcohol production is a common theme in food microbiology, using fermentation pathways in yeasts that yield ethanol. Cyanobacteria may potentially be used to produce metabolites such as ethanol by cloning fermentation pathways into cyanobacteria and producing these metabolites using energy from photosynthesis. Isobutanol production is also promising in cyanobacteria, with efficiency and yields potentially higher than that of ethanol production. Furthermore, since isobutanol has a higher density and octane rating relative to ethanol, isobutanol may be a better component of biofuel for societal use[9].

Production of hydrogen gas may be tapped by the use of hydrogenases, which catalyses the reversible reaction $2H^+ + 2e^- \rightarrow H_2$. Several cyanobacteria are also capable of evolving hydrogen gas under anoxic conditions, and the direct use of reducing equivalents from the light reactions of photosynthesis allows it to bypass the otherwise inefficient Calvin Cycle. However, light to hydrogen conversion is very inefficient in wild-type cyanobacteria and substantial improvements must be made for cyanobacterial hydrogen to be a viable fuel source [9].

Other photosynthates that are valued in industry includes isoprene and sugars. Isoprene is a volatile precursor for many compounds and can be polymerised by chemical processes to generate synthetic rubber. Isoprene is produced in many plant plastids, but isolation of isoprene from plants is not practical. Sugar production in cyanobacteria is the result of natural fixation of carbon, as it is in plants as well. Although artificial sugars have been synthesised in the lab, sugar production in cyanobacteria is more cost-effective. Production of isoprene and sugar may be possible in cyanobacteria, but improvements must be made to this expression system [9]

Hormogonium Development

Hormogonia are non-dividing motile filaments which seek to establish new colonies, including in environments with better sources of light, and are required for establishing symbiosis with new plant partners. Development of hormogonia results in filamentous cyanobacteria capable of gliding across surfaces [11]; and while the mechanism of motility has not been fully described, progress has been made in understanding the mechanism of motility in hormogonia, which is driven by a type IV pilus (T4P) system. During hormogonium development, hormogonia can be distinguished from vegetative filaments morphologically. Hormogonia filaments tend to be shorter than vegetative filaments, and cell sizes are smaller than those of vegetative cells, which is the result of reductive cell division upon induction. The cells at the end of the filaments become more tapered [12], presumably to assist in motility. Heterocysts are not present during hormogonia development, and therefore hormogonia are unable to fix nitrogen. During hormogonium development, cell division is arrested so that no biomass is accumulated and will resume upon reverting to the vegetative growth state [4]. Following induction, hormogonia remain motile for up to 48-72 hours [4], allowing the filaments to seek more favourable environments, including better sources of light, as well as to seek new symbiotic partners [12,13]. After finding a more favourable environment, they transition back to the vegetative growth stage [4], where they then establish a new colony and resume normal microbial metabolism, fixation of nitrogen in its absence, and other metabolic processes of vegetative filamentous cyanobacteria.

Numerous components are responsible for motility in cyanobacteria. Based on developmental studies using the filamentous cyanobacteria, *Nostoc punctiforme*, filamentous cyanobacteria capable of motility employ a Type IV Pili system and secrete a hormogonium

polysaccharide (HPS) which is thought to facilitate gliding motility. Two chemotaxis-like systems, the *Hmp* and *Ptx* systems, are activated in hormogonia, with the latter controlling phototaxis and the former essential for motility. All these components work together solely in hormogonia, and the genes coding for proteins involved in these systems have been found to be expressed mainly during this developmental stage.

Hormogonium development is a prerequisite for establishing plant-cyanobacterial symbioses [12], and it is believed that following colonization, plants provide sugars that inhibit further hormogonium development in order to maintain symbioses with cyanobacteria [14,15]. By fully understanding the relationships between plants and cyanobacteria on a phenotypic and genetic level, insights on hormogonium development may also lead to the artificial engineering of nitrogen-fixing symbioses with essential crop plants. These symbioses can reduce the need for a chemical fertiliser containing nitrogen, which may allow farmers to save money and allocate funds, which would otherwise have been spent on fertiliser, on other investments [10]. Additionally, the reduction of use of such fertiliser can reduce the runoff of nitrogen that lead to algal blooms[16].

Type IV pili mediate motility in filamentous cyanobacteria. Motility in cyanobacteria is generally performed by a Type IV Pilus (T4P) system (Figure 1). The T4P machinery varies between different bacteria, but always consists of a pilus that extends through a channel on the cell surface. As the pilus extends out of the cell, the pilus engages in motility by adhering to the surface that the filaments glide on. Once attached, the pilus is retracted, pulling the cell forward. The extension and retraction of the pilus is done through polymerisation and de-polymerisation of the pilus by motor proteins using pilins as building blocks. In *N. punctiforme*, PilA pre-pilins are used as building blocks for the pilus. PilA pre-pilins are digested into PilA pilins by PilD

and are assembled into the pilus by PilC. PilB and PilT ATPases mediate assembly and disassembly of the pilus respectively, which leads to extension and retraction. The pilus is extended and retracted through PilQ, which is a transmembrane protein sitting on the outer membrane and acts as a channel that guides the pilus out of the cell. These proteins all contribute to the formation of the pilus which allows for the filament to move.

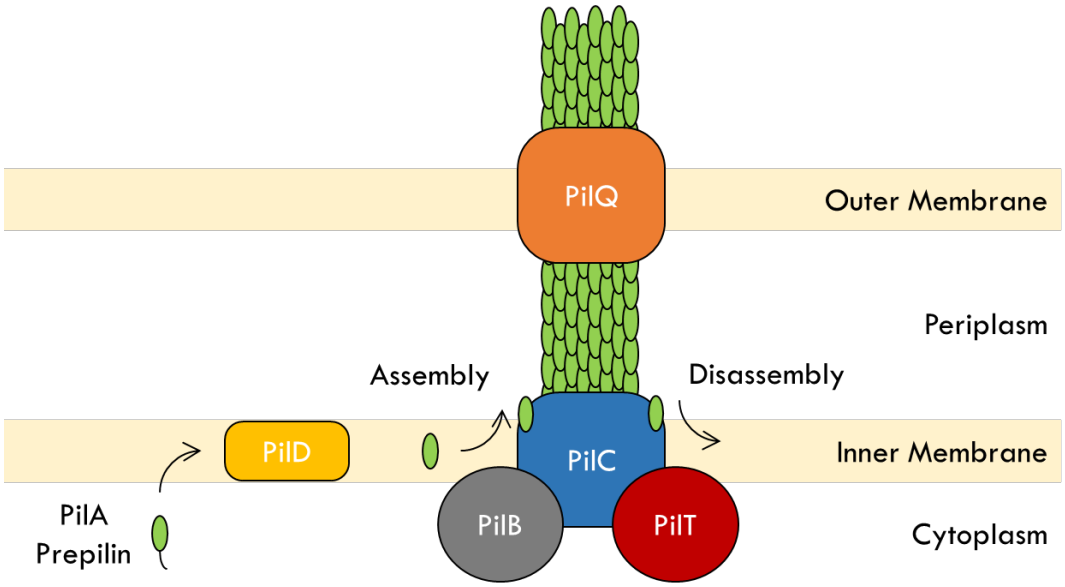

Figure 1: The type IV pilus (T4P) secretion system which mediates motility in hormogonia. The T4P is composed of PilA monomers. The PilA pre-pilins are processed by the protease, PilD, and is assembled into the pilus by PilC. PilB and PilT are motor proteins that interact with PilC to drive assembly and disassembly of the pilus respectively. Finally, the pilus extends out of the cell through PilQ, a channel protein on the outer membrane that facilitates passage of the pilus out of the cell. This simplified figure of the T4P was drawn based off Figure 2A from Hospenthal et-al 2017 [17].

Along with T4P, we see increased levels of polysaccharide secretion during hormogonium development and motility. This polysaccharide, synthesized by proteins encoded at the *hps* locus, is thought to facilitate gliding motility of the cyanobacteria along the surface. The *hps* locus encodes several putative glycosyl transferases and pseudopilins/minor pilins. We

are not certain of how polysaccharide contributes to motility, but previous studies have shown that knocking out polysaccharide secretion retards the movement of cyanobacteria. However, cultivation of *hps*-deficient mutants in a medium containing polysaccharide produced by wild-type *N. punctiforme* led to restored motility of these mutants, showing that this extracellular polysaccharide plays a role in motility [11,18].

Hormogonium gene regulatory network. Hormogonium development is regulated by a few different signal transduction systems: two chemotaxis-like systems, the Hmp and Ptx systems, and a partner switching regulatory system. These signal transduction systems are required for normal motility of hormogonia.

The Hmp chemotaxis-like system is required for hormogonium development. The Hmp chemotaxis-like system functions in regulating directional motility of cyanobacteria, but this system has only been recently discovered and has not been fully characterised. In *N. punctiforme*, these genes are encoded by the *hmp* locus, which includes *hmpABCDE*, and the divergently transcribed *hmpF* gene. The *hmp* genes encode chemotaxis-like proteins whose inactivation abolishes motility[19,20]. HmpF was found to be localised at the leading poles of the cells in motile hormogonia, indicating a possible role in activating the T4P motors, but the mechanisms regulating HmpF localization and motor activation are currently not clear[20].

The Ptx chemotaxis-like system appears to mediate phototaxis. *N. punctiforme* hormogonium exhibit phototactic behaviour in which filaments respond by migrating toward white light[13]. This phototactic response is thought to be mediated by the chemotaxis-like Ptx system, encoded at the *ptx* locus, which comprises the genes *ptxA-H*. The Ptx system includes the gene products of *ptxD*, encoding an MCP homologue, and *ptxE*, encoding a CheA homologue, in which inactivation of either did not abolish motility, but abolished net movement

toward light. In bacteria, these gene products make up a signal transduction system allowing it to respond to extracellular signals[21]. In the case of cyanobacteria, these homologues facilitates positive phototaxis.

A partner switching system is thought to regulate hormogonium development. Partner switching regulatory systems control gene expression generally by regulating the activity of sigma factors, which are involved in transcription initiation. In *Nostoc punctiforme,* we recently identified an Rsb-like partner switching regulatory system which is implicated in hormogonium development. In *Riley, Gonzalez, Risser 2018*[22], we described this partner switching system and determined that the proteins HmpU, a putative serine phosphatase, and HmpV, a STAS domain protein, enhance motility and persistence of the motility state, while HmpW, a putative serine kinase/anti-sigma factor, does the opposite. Rather than regulation of a sigma factor by HmpW, as observed in other Rsb-like systems, it appears that HmpV is the output of the partner switching regulatory system, which is active in its unphosphorylated state, when the phosphate is removed by HmpU, and in turn positively regulates an unknown downstream component to promote hormogonium development. HmpW, on the other hand, adds a phosphate group to HmpV and inactivates it.

Regulation of Gene Expression in Bacteria

All bacterial genes need to be regulated in some way as it allows the cell to determine which genes should be expressed. Because the genome of organisms often contains thousands of different genes, it is detrimental to the organism, and a waste of energy and resources, if all genes were to be expressed constitutively. A quintessential example of gene regulation in bacteria involves the *lac* operon, in which expression of these genes are only activated in the presence of lactose. By regulating which genes are expressed, cells can respond to internal and

external stimuli in order for the cell to survive the continuously changing environmental conditions.

Gene expression is regulated at four levels: transcription, post-transcription, translation, and post-translation. In this study, we are mainly interested in the transcriptional regulation of genes that are involved in hormogonium development. In order to ask questions regarding transcriptional regulation, we first need to understand how DNA is transcribed in bacteria.

Genes in bacteria are often arranged into clusters of similar genes known as operons. An operon generally consists of a promoter, an operator, and several genes encoding translational products. The promoter is a region where an RNA polymerase (RNAP) binds to and initiates transcription, and consists of two regions, the -10 and -35 regions, which RNAP associates with during initiation of transcription. An operator usually functions in regulation of the operon, usually via a repressor that binds to the operator and physically blocks RNAP from associating with the promoter or performing its transcriptional activity. When the gene needs to be expressed, the repressor would dissociate in some way in order to allow transcription to occur.

Sigma factors and their role in transcription. Bacterial RNAPs require the presence of certain transcription factors in order to transcribe all genes in the bacterial genome. The activity of RNA polymerase (RNAP) core enzyme is regulated by these transcription factors and its affinity to certain promotor regions depends on their presence. These transcription factors, which bind directly to the RNAP core enzyme, are called sigma subunits, or sigma factors, and produce an RNAP holoenzyme which allows it to target promoter DNA for transcription. The eubacterial RNAP holoenzyme, which is fully capable of binding to and transcribing genes, is composed of the σ subunit and the core enzyme, which is composed of α_2 (RpoA $\times$ 2), β (RpoB), β′(RpoC) and ω subunits[23,24]. The core enzyme contains the catalytic activity

needed to read and transcribe DNA into messenger RNA. However, it does not possess any helicase activity, nor can it bind to specific sequences or motifs in the promoter region; that is, the polymerase does not know where to initiate transcription. The σ subunit possesses such activity and is required for initiation of transcription. It associates with the RNAP core enzyme, forming an RNAP holoenzyme and the holoenzyme can now bind to a promoter and initiate transcription[25]. It specifically looks for the -10 and -35 regions of the promoter DNA and binds to it. Once bound to the promoter, the sigma subunit initiates strand separation through a mechanism that is not fully characterised, but it is believed that a few nucleotides of the AT-rich -10 region are flipped out of position into the binding pockets of the sigma subunit, weakening the hydrogen bonds between the double stranded DNA so that the strands are separated, allowing RNAP to initiate transcription[24]. Once transcription has begun and a few ribonucleotides have been incorporated into the mRNA, the σ factor dissociates from the core enzyme and may bind to another RNAP to initiate transcription of another operon [25].

There are several different sigma factors in bacteria, but they generally consist of the same features. Four distinct domains generally make up a sigma factor: the σ1.1, σ2, σ3, and σ4. The σ1.1 domain interacts with the rest of the sigma factor domains and prevents the sigma factor from interacting with DNA in the absence of RNAP. The σ2 domain interacts with DNA and the RNAP core enzyme. This domain can be divided into several distinct sub-domains: σ1.2, σ2.1, σ2.2, σ2.3, and σ2.4. The σ1.2 sub-domain interacts with a discriminator element on the DNA, which enhances transcription of its promoter during the stringent response, where (p)ppGpp is produced upon nutrient limitation and binds to RNAP[26]. The σ2.1 and σ2.2 sub-domains interact with the β' subunit of RNAP core enzyme. Finally, the σ2.3 and σ2.4 sub-domains associate with the -10 region of the promoter DNA. The σ3 domain interacts with an

extended -10 region, which if present, usually helps stabilise the initiation complex. Finally, the σ4 domain interacts with the -35 region of the promoter via a helix-turn-helix domain. It also binds to the β subunit of RNAP and any transcriptional activators upstream of the promoter.

In addition to acting as a specificity factor for transcription initiation, sigma factors also play an important role in the opening of double-stranded DNA[24]. DNA melting occurs in the -10 region, which consists primarily of adenosine and thymine nucleotides, which form two hydrogen bonds for each base pair. The number of hydrogen bonds compared with guanines and cysteine base pairs helps facilitate DNA melting, which is done by the sigma factor. The mechanism of this process is not well understood, but crystallisation studies are giving more insight. A recent study where the interaction between a sigma 2 subunit and the -10 sequence was crystallised shows two nucleotides interacting with binding pockets on the subunit. These nucleotides were flipped in order to interact with their respective pockets, which facilitates in the opening of the double stranded DNA.

Alternative sigma factors. Sigma factors act as a specificity factor, where each different sigma factor allows the RNAP to target certain groups of genes based on their promoter sequences[25]. Since bacteria express a number of different sigma factors, we can infer that each different sigma factor can initiate transcription of operons with a certain promoter sequence. Sigma factors in bacteria can be subdivided into four groups. Group 1 sigma factors are generally essential for cell viability. They are known as housekeeping sigma factors and are involved in the expression of housekeeping genes. Any disruption of expression of these housekeeping genes will compromise the viability of the cell and usually lead to the death of the cell. The housekeeping genes are involved in keeping essential pathways operating including critical metabolic pathways and the removal of wastes.

Group 2 sigma factors are structurally similar to group 1 but are non-essential for cell viability. These sigma factors are responsive to changes in the environment and are generally activated during the changes in light quantity, periods of heat shock, dehydration, or nitrogen-starvation. Group 3 sigma factors are slightly homologous to groups 1 and 2 in their amino acid sequences and are involved in the transcription of regulons for survival under stress due to changes in environmental and developmental conditions[25]. Group 4 sigma factors, which are known as extracytoplasmic function (ECF) sigma factors, are involved in response to cell envelope stress. These sigma factors are inactive and bound to an anti-sigma factor which resides as a transmembrane protein. Some extracellular stressor would cause the proteolysis of the anti-sigma factor, releasing the ECF sigma factor, which can activate genes involved with cell envelope stress[27].

Alternative sigma factors in *N. punctiforme*. In addition to its housekeeping sigma factor, *N. punctiforme* has 12 alternative sigma factors which can be organised into the four groups based on their structure (Table 1). The group 1 sigma factor SigA acts as the housekeeping sigma factor. The group 2 sigma factors have been described to be involved in heterocyst differentiation[28] while group 3 sigma factors have been described to be involved in desiccation tolerance[29], both in *Anabaena/Nostoc* PCC 7120. Group 4 sigma factors are also known as the ECF sigma factors. In *N. punctiforme,* there are two potential ECF sigma factors, SigG and SigK, with SigG recently confirmed as an ECF sigma factor with its regulon identified[30].

Table 1: Sigma Factors found in *N. punctiforme*

Sigma Factor	Group
SigA	Group 1
SigB-a, SigB-b, SigB-c, SigB2, SigB4, SigC, SigD, SigE	Group 2
SigF, SigJ	Group 3
SigG, SigK	Group 4

Alternative sigma factors in cyanobacteria have been extensively studied, including sigma factors which are the subject of this report. However, many studies on sigma factors in cyanobacteria have been conducted in unicellular cyanobacteria and not filamentous cyanobacteria, and the results of such studies do not bear great significance on their effects in development and cell differentiation, a phenomenon only occurring in filamentous cyanobacteria. Sigma factor studies that are conducted on filamentous cyanobacteria are generally studies involving heterocyst development, but to date, no sigma factor has been shown to definitely control development.

Previous work on SigC, SigF, and SigJ. Sigma factor C (SigC) is a group 2 sigma factor that was found to be involved in heterocyst differentiation[25,28,31]. Prior studies have shown that *sigC* is activated whenever nitrogen is scarce, and that inactivation of *sigC* resulted in downregulation of 58 genes at 8 hours after nitrogen deprivation in the filamentous cyanobacterium *Anabaena*. *sigC* was also found to be upregulated in the heterocysts themselves, along with *sigE*, another group 2 sigma factor, and *sigG*, a group 4 sigma factor[28,30]. Heterocyst development in a *sigC* mutant was delayed, but not abolished, possibly because *sigE* may have partially overlapping promoter specificity with *sigC*. Therefore, SigC helps promote enhanced heterocyst development, but is ultimately not required[28].

From previous studies with the cyanobacteria *Synechocystis* sp. strain PCC 6803, the group 3 sigma factor, *sigF*, was found to be involved in phototactic movement and pili formation in cyanobacteria [32]. Studies using *sigF* mutants showed that *sigF* is crucial for the motility of cells, as *sigF* mutants lost their phototactic motility, as well as their pili, which are abundant on the surfaces of wild-type cells.

The group 3 sigma factor, *sigJ*, has been reported to be involved in the regulation of genes involved in desiccation tolerance, and is activated during a period of dehydration of the cell. In *Nostoc* PCC 7120, higher expression of *sigJ* resulted in the upregulation of 112 genes, including *sigJ* itself and the downregulation of 42 genes[29].

While these sigma factors have been studied mainly during the development of heterocysts, it is unclear how deletion of each sigma factor affects gene expression in hormogonia.

Deletion of *sigC*, *sigF*, and *sigJ* in *N. punctiforme* and Experimental Plan

As previously mentioned, countless studies have been conducted on filamentous cyanobacteria to better understand the gene regulatory networks regarding development. A popular platform for studying development of filamentous cyanobacteria is *Anabaena/Nostoc* sp. strain PCC 7120. However, this laboratory strain is incapable of producing hormogonia, which is probably attributable to the accumulation of mutations through prolonged culturing[20], producing results that may not fully describe the regulons of these sigma factors. Therefore, we intend to use *N. punctiforme* ATCC 29133, which is capable of differentiating into hormogonia, allowing us to study and identify the role of these sigma factors in hormogonium development.

Deletion of each sigma factor gene abolished motility, but each deletion produced a different phenotype (Figure 2). Deletion of *sigC* or *sigJ* produced strains incapable of

differentiating into morphologically distinct hormogonia, and both failed to accumulate extracellular PilA, a component of the type IV pilus, or HPS (hormogonium polysaccharide). The Δ*sigC* strain was nevertheless able to produce wild-type levels of intracellular PilA while PilA levels in the Δ*sigJ* strain were undetectable. Additionally, the cell size of the Δ*sigC* strain was also affected, with the vegetative cells of the Δ*sigC* strain (4.39+/-0.5 μm) being larger than those of the wild-type (3.41+/-0.47 μm), indicating that SigC may be involved in cell division and may be required for production of hormogonium morphology. The Δ*sigF* strain is capable of differentiating into morphologically distinct hormogonium but failed to produce detectable levels of both intracellular and extracellular PilA like in Δ*sigJ*.

The differences in phenotypes of the sigma factor mutants implies that *sigJ* may be expressed early in hormogonium development, and its expression may also control *sigC* and *sigF.* Expression of *sigF* appears to occur later in hormogonium development and its expression is required for *pilA* expression while *sigC* expression may be required for expression of genes involved in cell division and morphology.

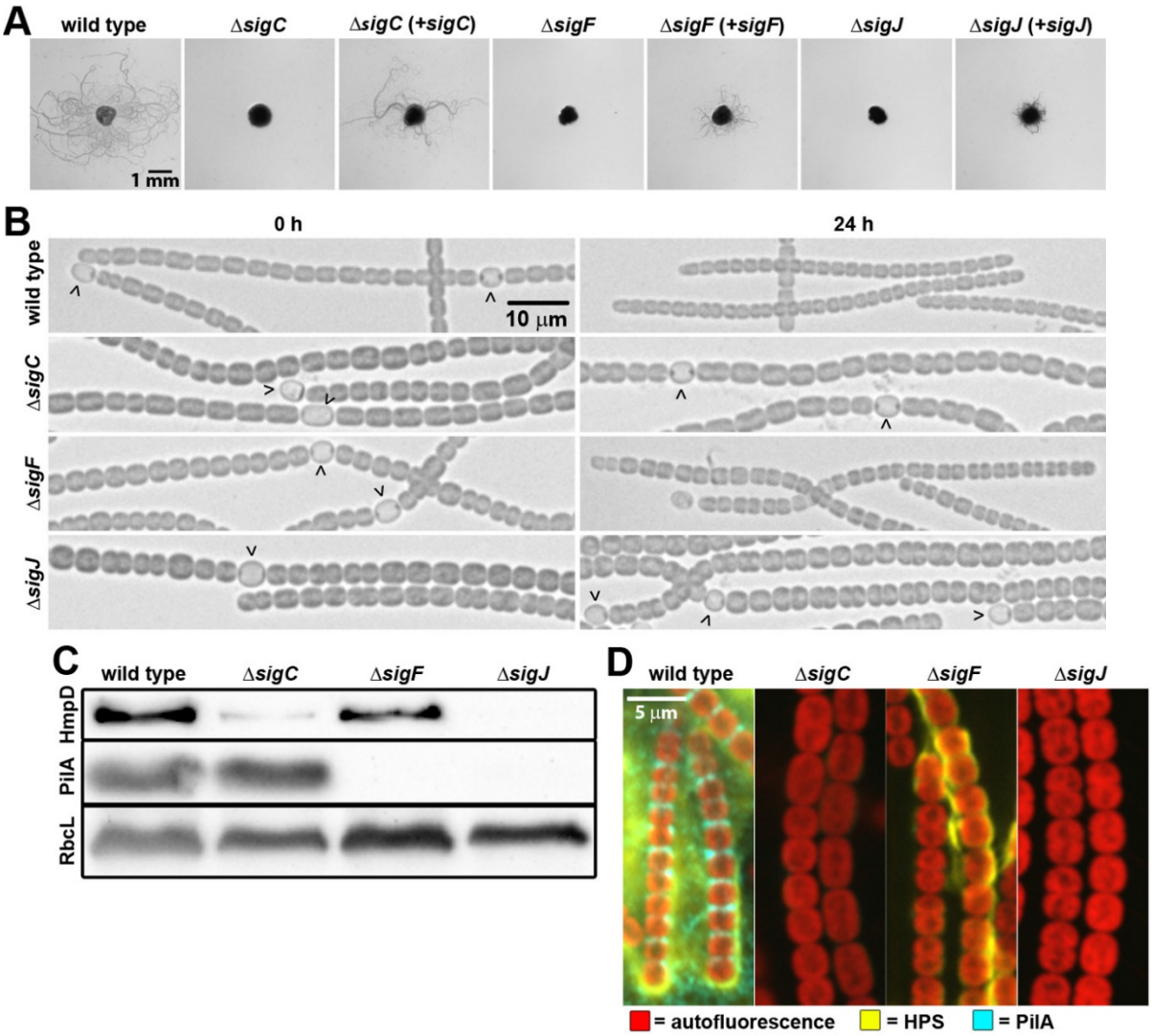

Figure 2: Phenotypic analysis of sigma factor deletion strains. (A) Plate motility assays performed. Plate motility assays done on sigma factor deletion strains showed a loss of motility, which was restored upon introduction of the sigma factor gene via a shuttle vector. (B) Light microscopy images of wild type and sigma factor deletion strains of *N. punctiforme* before and 24 hours after induction of hormogonium development, with heterocysts indicated with an arrow. (C) Immunoblot analysis of HmpD, and PilA in the wild-type and deletion strains of *N. punctiforme* 24 hours after hormogonium induction. RbcL, the large subunit of RuBisCO, served as a loading control. (D) Immunofluorescence and lectin staining analysis of extracellular PilA and HPS. Depicted are merged images of fluorescence micrographs acquired using a 63X objective lens from cellular autofluorescence (red), PilA immunofluorescence (cyan), and UEA-fluorescein-stained HPS (yellow) for various strains 24 h after hormogonium induction.

Therefore, the goal of this study was to assess expression levels of several genes involved in hormogonium development in each sigma factor deletion mutant using real-time PCR and RNA-seq. Because knockouts of *sigC*, *sigF*, and *sigJ*, led to defective hormogonium filaments,

we predicted that many genes involved in hormogonium development will have been down-regulated as a result. We tested the effects of these deletions using real-time PCR with primers for 20 genes previously identified as important for hormogonium development, and RNAseq was utilised to confirm the RT-qPCR data and to further describe the regulons of these sigma factors. The results of this study may be used as an addendum to the current model of the gene regulatory network of hormogonium development and further contribute to the understanding of development in filamentous cyanobacteria.

Chapter 2: Materials and Methods

Culturing of *Nostoc punctiforme*

Allen and Arnon diluted four-fold (AA/4) medium was prepared by mixture of 3.15ml of -Pi and 1.6ml of +Pi into 500ml of double-distilled water. The medium was autoclaved at 121°C for 30 minutes. *N. punctiforme* strains (Table 1) were cultured in 50ml of AA/4 medium with 4mM sucralose, which inhibits hormogonium development. Cultures were grown under constant illumination at room temperature on a shaker at 125 RPM. Cultures were incubated for ~2 weeks. After incubation, 25mL of the 50mL culture were transferred to 500ml of AA/4 with 5mM sucralose. 5mL of the 50mL culture were transferred to a new 50ml culture for growth of subsequent replicates. The 500ml cultures were grown for 2-3 more weeks under the same conditions until chlorophyll-*a* concentration of ~3.5µg/ml was reached. Three biological replicates were grown.

Quantification of Cell Material Using Chlorophyll-a Concentrations

For each culture, 1ml of cells were taken and centrifuged for 3 minutes at 2,000 RCF. 900 µl of supernatant were replaced with 100% methanol and the cells were re-suspended via vortexing. The reaction mixture was incubated for ~5-10 minutes in the dark and then centrifuged for 1 minute at maximum speed. The optical density of the supernatant was taken by a Denville Scientific S1200 Diode Array Spectrophotometer by reading the absorbance at 665nm. Chlorophyll-*a* concentrations and total volumes of culture needed for 300 µg of Chl*a* were calculated using equation 1 and equation 2 respectively.

(1) $$\mathrm{OD}_{665} \times 12.7 \frac{\mu g}{ml}$$

(2) $$\text{chlorophyll-}a\text{ concentration}\,\frac{\mu g}{ml} \times (300\text{ng cells})^{-1}$$

Hormogonium Induction and Sample Collection

Induction of hormogonium development in *N. punctiforme* was done by the removal of sucralose from its growth medium. For each strain, the equivalent of 1,200 µg of Chl*a* in cell material of *N. punctiforme* to be induced were taken from the 500ml culture. The medium was removed by centrifugation for 5 minutes at 2,000 G and 24°C in 50ml conicals, the cells washed with 50ml AA/4 without sucralose, centrifuged again, and the supernatant decanted. This was repeated a 2nd time. The harvested cells were then re-suspended in 40ml of AA/4 without sucralose and then divided into four flasks containing 40mL of AA/4, so that the equivalent of 300 µg Chla of cell material was in a final volume of 50ml. The induced cells were incubated under constant illumination on a shaker at 125 RPM at room temperature. Induction of hormogonium development was allowed to occur for up to 18 hours, with one of each of the four flasks harvested at 1, 6, 12, and at 18 hours respectively. Each harvested 50ml culture was centrifuged for 5 minutes at 2,000 G at 24°C, and had most of its growth medium removed, leaving around 2ml. The cells were re-suspended and transferred into 2ml RNAse-free screw-cap tubes (GeneMate) and centrifuged for 2 minutes at maximum speed. Medium was removed, only leaving 500mL, and the tubes were immediately flash-frozen in liquid nitrogen for 5 minutes. The tubes were kept in storage at -80°C until RNA extraction. An equivalent amount of un-washed, un-induced cells, which served as the 0-hour time point, were collected, concentrated to 500ml without washing, and immediately flash-frozen using liquid nitrogen. Three biological replicates were collected.

RNA Extraction, DNAse Treatment, and Dilution

For lysis, 0.58 g of 0.5 mm zirconium/silicate beads (Biospec), 33.3 µl of RNAse-free water, 167 µl of 3% Celite, and 583 µl of Tris-buffered phenol were added to the still-frozen cells and shaken in a MiniBead Beater (Biospec) for 160 seconds at high speed and immediately placed on ice. The lysed cells were spun for 15 minutes at 4°C at maximum speed. 650µl of the top layer (aqueous) was transferred to new tubes containing 600 µl chloroform and rigorously mixed. The mixture was spun down for 10 minutes at 4°C at maximum speed. 500µl of the aqueous (top) layer was transferred to new tubes containing 182 µl precipitation solution (7.5 M lithium chloride, 20 mM Tris, pH 7.4, 10 mM EDTA, pH 8), and mixed by inversion. The RNA was precipitated at -80°C for 1 hour, and then spun for 15 minutes at 4°C. The supernatant was discarded, and the pellet was washed and re-suspended with 1ml of 70% ethanol. The re-suspended RNA was centrifuged for 10 minutes at 4 °C at maximum speed. The ethanol was removed, and the wet RNA pellet was re-suspended in 100µl of TE buffer.

For removal of chromosomal DNA, DNAseI was used in conjunction with the RNeasy Mini Kit (Qiagen). The clean-up was conducted in accordance with manufacturer instructions. Before the 2nd wash with Buffer RW1, the RNA was treated with DNaseI (catalogue no. #79254, Qiagen) in accordance with manufacturer instructions for 15 minutes. The RNA was quantified using the NanoDrop 2000 (ThermoFisher Scientific), using RNAse-free water as a blank. The RNA was subsequently diluted to 500ng/µl.

cDNA Synthesis for RT-qPCR

cDNA was synthesised using the ProtoScript First Strand cDNA Synthesis Kit (New England Biolabs) using random hexamers according to manufacturer instructions. 80µl of autoclaved water was added to the reverse transcription product to yield a total of 100µl of cDNA.

Real Time PCR and Analysis

Genes selected for RT-qPCR and RNAseq analysis are genes shown to be activated in hormogonium development (Table 2[4,11,13,20,22,33]). Forward and reverse primers were designed based on the sequence of the genes using Primer3 v0.4.0 at default settings, in which the gene sequences were taken from the cyanobacteria genomic database, CyanoBase[34]. Primers were ordered from Integrated DNA Technologies, and came in a lyophilised stock, which was diluted to 50µM. Subsequently the forward and reverse primer pairs were mixed (final concentration 10µM each). The primers were tested using *Nostoc punctiforme* ATCC 29133 genomic DNA. Each qPCR used 1.6µl of primer mix, 6.4µl of water, and 10µl of SensiFAST SYBR No-ROX Real Time PCR mix (Bioline). 2µl of the chromosomal DNA was added to each reaction. qPCR was conducted using a StepOnePlus Real Time PCR System (ThermoFisher Scientific). The initial denaturation step first ran for 20 seconds at 95°C. Afterwards, 40 cycles of qPCR were run, which consisted of a 3 second 95.0°C denaturation step, a 10 second 62.0°C annealing step, and a 10 second 72.0°C elongation step. A melt curve analysis was performed to ensure a pure PCR product with an initial 15 second 95.0°C denaturation step followed by a 1 minute 60.0°C renaturing step, and then 0.3°C incremental increases. Primer efficiencies for all primer pairs were quantified based on a dilution series of chromosomal DNA and were all found to be above 90%. For analysis of gene expression, the

same run method was used, using the appropriate primer mix and cDNA for each reaction. The melt curve analysis was performed as described earlier but with 1.0°C incremental increases. Data from the RT-qPCR were exported into a Microsoft Excel spreadsheet using the StepOnePlus software. RT-qPCR data were normalized using the $2^{-\Delta\Delta Ct}$ method[35] using *rnpB* (Npun_R018) expression as a reference. A log2 transformation was performed and the triplicates were averaged. Heat maps were generated using Genesis[36].

RNA-seq and Analysis

For transcript mapping using RNA-seq, library preparation and sequencing was carried out at the University of California Berkeley QB3 Vincent J. Coates Genomics Sequencing Laboratory using 10 µg of total RNA extracted from all strains 0, 1, 6, 12, and 18 hours post-induction. Ribosomal RNA was removed using the Ribo-Zero rRNA Removal Kit (Bacteria) (Illumina, Inc.). Directional cDNA libraries were synthesized from the rRNA depleted samples, sheared to a library size of 200 bp, appended with adapters, and sequenced across 6 lanes of an Illumina HiSeq4000 flow cell, multiplexed with 4 other libraries, generating 100 bp paired end reads. Alignment, assembly, normalisation, and quantification of sequencing data was performed using the software package Rockhopper[37,38] using default parameters and transcript maps were generated using Integrated Genomics Viewer[39]. RNA-seq data was deposited in the NCBI GEO database (GSE124969).

Differential expression for each strain, time point, and replicate was calculated as log2(normalized expression of experimental condition/normalized expression of wild-type strain at T = 0 h [mean of 3 biological replicates]). Statistical identification of differentially expressed genes in the wild-type time course and between the time course for the wild-type and sigma factor deletion strains was performed using Bayesian Analysis of Time Series (BATS) using

default parameters[40]. Differentially expressed genes were subsequently binned into one of four categories. The first contained genes that were differentially expressed during wild-type hormogonium development and whose expression was not statistically different in any of the sigma factor deletion strains. The remaining 3 categories contained genes that were differentially expressed in the wild type but showed an altered expression pattern in one or more of the sigma factor deletion strains. These genes were assigned to one of three bins, one for each sigma factor, based on which deletion strain produced the largest L2-norm value. Each bin was subsequently subjected to a hierarchical cluster analysis (unweighted pair group method using average linkages [UPGMA]) using the software package Genesis[36]. Heat maps depicting expression data were generated in Genesis[36]. Groups of 5 or more genes with similar expression patterns that were distinct from their neighbors were then manually assigned a cluster number.

Table 2. Strains used in the study

Strains	Relevant Characteristic(s)	Source
Nostoc punctiforme strains		
ATCC 29133	wild type	ATCC
UOP131	Δ*sigC* (Npun_F0996) *	This study
UOP132	Δ*sigJ* (Npun_R1337)	This study
UOP142	Δ*sigF* (Npun_F4811)	This study

* locus tag denoted in parentheses

Table 2: List of strains of *Nostoc punctiforme* used in this study. The ATCC 29133 strain was obtained from ATCC. The sigma factor deletion strains were previously made using site-directed mutagenesis.

Table 3. Genes and primers used in the study

Gene	Locus Tag	Direction	Primer Sequence
rnpB	Npun_r018	Forward	TAAGAGCGCACCAGCAGTAT
		Reverse	CATTGAGCGGAACTGGTAAA
ogtA	Npun_F0677	Forward	TGCGGCAAAATAAAGTAGCA
		Reverse	TCCTCGCTGGATTAGGTGTT
pilA	Npun_F0676	Forward	TCTGGTTGCCAACAATGGTA
		Reverse	ACTTCAGCACTCCGATCACC
sigC	Npun_F0996	Forward	GGCGATCGCAACTTCTAGTC
		Reverse	ACTTGGGTCGGTGTCATCTC
sigF	Npun_F4811	Forward	TGTTTGGGAGAATTGGTTCC
		Reverse	ATCCCGAGATGTTCTGCAAC
sigJ	Npun_R1337	Forward	TGAGATGCTGCACTTTTTGC
		Reverse	TTTTGAGCGGCTAACTTGGT
hmpD	Npun_F5963	Forward	TAGTGATGCTTTGCGTCAGG
		Reverse	GCTTCACCACCTTTGAGAGC
hpsE	Npun_F0070	Forward	GGTAGCCAAATTCACCCTGA
		Reverse	TTGCCTTGAACTCTCCCAGT
pilB	Npun_R0118	Forward	AATGGTGTCGGCTACAAAGG
		Reverse	TCGGCTTCCAAACCAGTATC
hmpU	Npun_R5135	Forward	AAAAGTGCAAGCCTGTTGCT
		Reverse	CGGCGCAAAACTGAAATAAT
hpsA	Npun_F0066	Forward	TGTAATGCCGCAGCTACAAC
		Reverse	TGATCCCCTCCTGTAGCATC
pilQ	Npun_F5008	Forward	GCACCAATTCCCTGACCTTA
		Reverse	ATGCTGCACCACCATCATTA
hmpF	Npun_R5959	Forward	CAACAACTGGAGCAGCAAAA
		Reverse	CTGGAAATCCGCTTGATGTT
ftsE	Npun_F5138	Forward	CTGTTTATCACCGGGCCAAG
		Reverse	ACCCTTGAGCTTGCAGTACA
ftsI	Npun_F0168	Forward	TTCACCCACTCGCACTGTAT
		Reverse	CCTTGATGTAGCCGCCATTC
ftsZ	Npun_R4804	Forward	GCTCTCACTGTTGGGGTAGT
		Reverse	GGTGTTTGTTCGGGGATCAC
ftsQ	Npun_R4806	Forward	CCCAATCTTTGTGGCGGATT
		Reverse	TACCCTGATTTGCCCCACTT
mreC	Npun_R1840	Forward	GGTAGTGCTTGGGTATTGCG
		Reverse	CGTGATGCCAGTGGTTCTTT
ptxD	Npun_F2164	Forward	ATGGCTCCCCTGTTAAAGCT
		Reverse	GGCTGCTAACAAACCCCAAA

Table 3: A list of forward and reverse primers used in the study. Two sets of forward and reverse primers were designed and tested, in which the better performing set were used in the study.

Chapter 3: Results

A Hierarchal Sigma Factor Cascade Controls Hormogonium Development

Sigma factors form part of the RNAP holoenzyme and are required for targeting to the promoter in order to initiate transcription of its targets. From previous studies, it was found that sigma factors C, F, and J are essential for hormogonium development, and that deletion of any of these sigma factors abolished motility. While the Δ*sigF* strain was able to fully develop into hormogonia, albeit without motility, the other two sigma factor deletion strains, Δ*sigJ* and Δ*sigC*, showed hormogonium development that appeared to be arrested. In order to determine which genes were affected in the sigma factor deletion strains, RNAseq and RT-qPCR were used to measure transcription levels in developing hormogonia.

In the wild-type strain, *sigC* and *sigJ* were strongly activated immediately following induction (Figure 3A&C). However, *sigJ* showed additional activation later in the induction at around 12 hours. In contrast, the activation of *sigF* was slower, peaking around 12 hours (Figure 3B). When *sigJ* was deleted, expression of both *sigC* and *sigF* were both greatly reduced. Expression levels of *sigC* were ~64 times lower prior to induction when compared to the wild-type strain, and showed similar differences in expression levels after induction, indicating that SigJ may be required for expression of *sigC*. In contrast, expression levels of *sigF* before induction were similar to that of the wild type, but during induction, the *sigJ*-deletion strain showed a 1,000-fold reduction of *sigF* expression compared to the wild type at 18 hours, indicating that SigJ is required for expression of *sigF*.

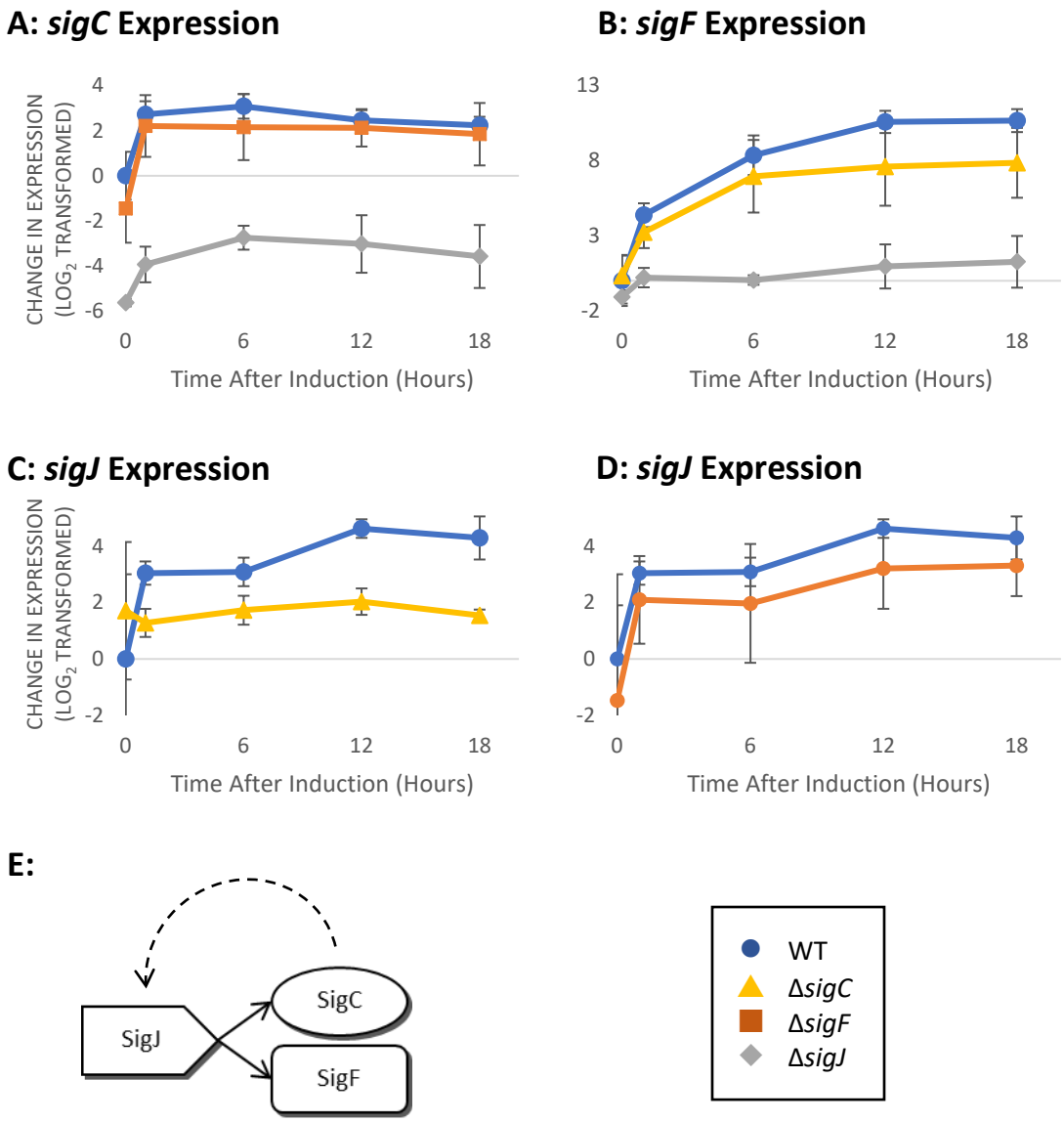

Figure 3: Relationship between the trio of sigma factors with error bars indicating standard deviations. Changes in expression of (A) sigma factor C, (B) sigma factor F, and (C & D) sigma factor J which were observed in ATCC-29133-WT (●), ATCC-29133-Δ*sigC* (▲), ATCC-29133-Δ*sigF* (■), and ATCC-29133-Δ*sigJ* (◆) strains of *N. punctiforme*. In the absence of SigJ, transcription of both *sigC* and *sigF* were greatly reduced, and therefore may be required for transcription of both sigma factors. A positive feedback loop may also be present, with SigC contributing to further activation of *sigJ*. (E) A summary of the relationships between the sigma factors.

Deletion of *sigC* reduced but did not abolish expression of *sigJ*. The reduced activation of *sigJ* in the absence of SigC may indicate the possible presence of a positive feedback loop, in

which SigC enhances activation of *sigJ*, and vice versa (Figure 3C). Expression of *sigF* in the Δ*sigC* strain was largely unaffected (Figure 3D). Deletion of *sigF* showed unremarkable results as expression of neither *sigC* nor *sigJ* were affected. These results suggest a hierarchal sigma factor cascade in which *sigJ* is activated by an unknown upstream protein, and SigJ activates *sigC* and *sigF*. Expression of *sigJ* may be enhanced by a positive feedback loop with *sigC* (Figure 3E).

Expression of Most T4P Genes is Dependent on Expression of *sigJ*, Except For *pilA*

Motility in hormogonia is driven by a type IV pilus system in which the pilus is used to pull the filament forward. Analysis of genes involved in activity of the T4P showed that the T4P motor protein gene, *pilB* (NpR0118), and the pilus channel protein gene, *pilQ* (NpF5008), both showed lower expression in the Δ*sigJ* mutant (Figure 4). However, although *pilA* was affected by deletion of *sigJ*, deletion of *sigF* completely abolished expression of *pilA*. Because *sigF* depends on expression of *sigJ*, *pilA* may ultimately be dependent on *sigF*. Deletion of *sigC* did not substantially affect expression of any of these genes during hormogonium development. Analysis of expression of these T4P genes indicate that their expression is dependent on *sigJ*, with the exception of *pilA* (NpF0676), which was ultimately strongly dependent on *sigF*.

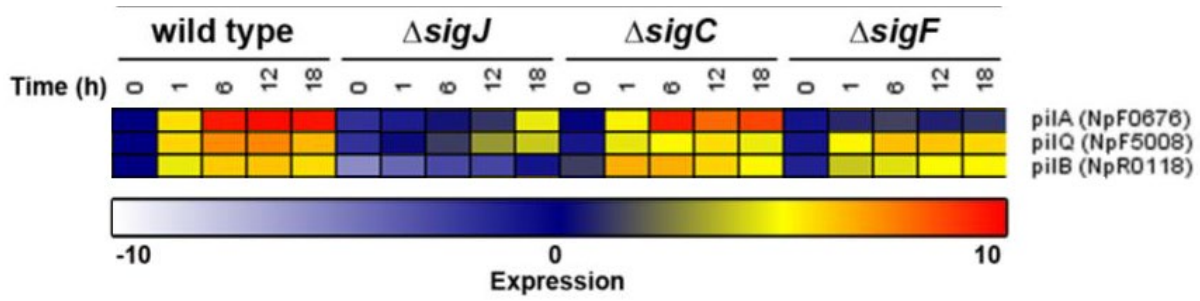

Figure 4: Expression of T4P genes. RT-qPCR data showing transcription levels of *pilA* in *N. punctiforme* and its mutants with differences in expression relative to the wild type shown as log2 transformed differences. The heat map shows expression differences of *pilA* (NpF0676), *pilQ* (NpF5008), and *pilB* (NpR0118) relative to ATCC-29133 before induction. Transcription of *pilA* was affected most in the Δ*sigF* mutant, but also in Δ*sigJ*, due to the absence of SigF. *pilB* and *pilQ* transcription levels were reduced only in the Δ*sigJ* mutant.

The Hmp Chemotaxis-Like System is Dependent on *sigJ*, Although not Stringently

The *hmp* locus produces gene products which are part of the Hmp chemotaxis-like system, and deletion of these genes abolished motility[19]; however, their exact functions and mechanisms are still not fully understood. *hmpD* (NpF5963) representing the *hmpBCDE operon*, and *hmpF* (NpR5959), which is divergently transcribed (Figure 5A) both showed immediate activation in the wild type, with ~64-fold increase in transcription, which peaked at 1 hour. Expression of these genes was largely unaffected in the Δ*sigC* or Δ*sigF* strains but was reduced in the Δ*sigJ* strain (Figure 5B). This indicates that although the *hmp* locus may be dependent on SigJ for robust activation, it is not stringently dependent on SigJ as some activation, especially at 18 hours, was still observed.

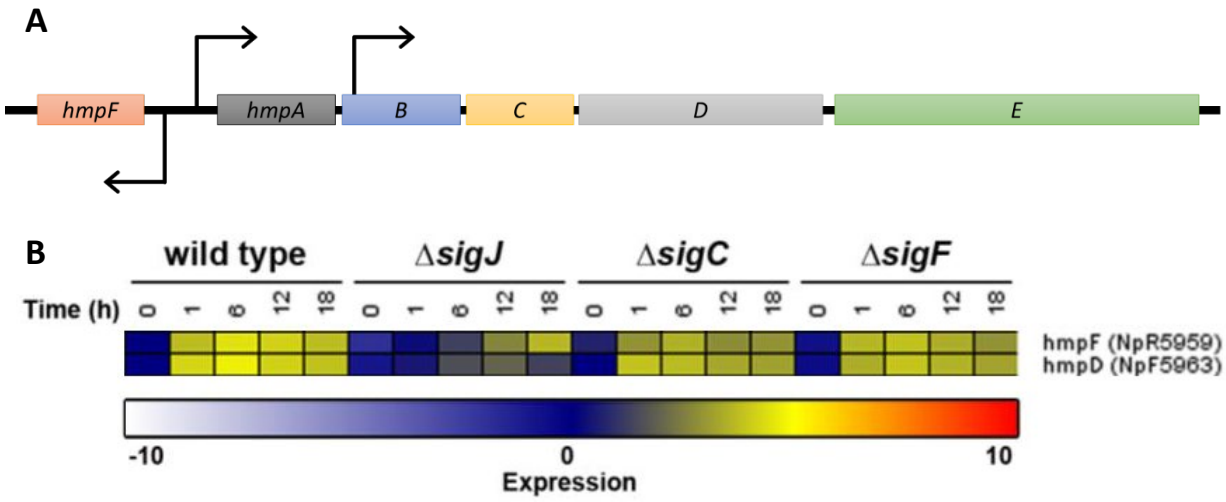

Figure 5: Expression of *hmp* genes. (A) The *hmp* locus. *hmpABCDE*, *hmpBCDE,* and *hmpF* are all under their own promoters. (B) Transcription levels of *hmpD* (NpF5963) and *hmpF* (NpR5959), as expressed in differences in transcription compared to the wild type strain at 0 hours after induction using $\log_2$ transformed values. Levels of transcription of *hmpD* (and the rest of the hmp operon) and *hmpF*, which is under a different promoter both saw reduction in transcription levels in the Δ*sigJ* mutant.

The Ptx Chemotaxis-Like System That Modulates Phototaxis Is Dependent on *sigJ*

Expression of *ptxD* was immediate in wild type hormogonia and peaked at 6 hours with a 200-fold increase in expression compared to pre-induction. The Δ*sigJ* mutant showed a 128-fold decrease in expression when compared to the wild type at 6 hours and increased slowly afterwards. Expression of *ptxD* was not drastically impacted in the other sigma factor deletion strains, however expression of *ptxD* in Δ*sigC* and Δ*sigF* was not as strong as in the wildtype around 6-12 hours (Figure 6). Therefore, the Ptx locus is dependent on sigJ for its expression, but *sigC* and *sigF* may play a role in enhancing the expression of genes within the Ptx locus.

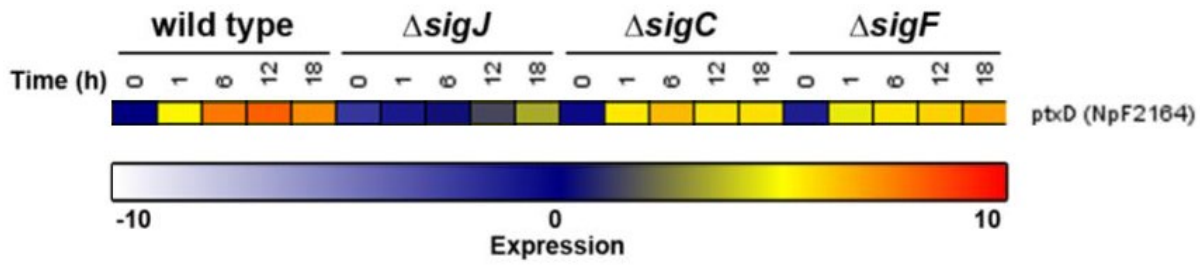

Figure 6: Expression of *ptxD*. Heat map showing expression levels of *ptxD* measured in changes of expression in $\log_2$ transformed values versus the wild type strain at 0 hours. Expression of *ptxD* was activated strongly in the wild type, *ΔsigC*, and *ΔsigF*, strains, but did not activate in *ΔsigJ*. A small amount of activation occurred at 18 hours in Δ*sigJ*, much later than what was normally expected.

HPS Production and Secretion Genes were Affected in Δ*sigJ* Mutants, but Still Showed Some Induction

Genes involved in polysaccharide production and secretion are essential for gliding motility in cyanobacteria. These genes are thus termed "hormogonium polysaccharide" or HPS genes and are encoded at the *hps* locus. While their functions are not fully understood, it has been shown that knockouts of these genes prevented production of detectable polysaccharide and abolished motility[41]. The two genes analyzed from this locus are *hpsA*, a gene that is conserved in many species of filamentous cyanobacteria and is likely involved in HPS secretion, and *hpsE*, which encodes a glycosyl transferase. Activation of *hpsA* was gradual, eventually peaking at 12 hours in the wild type (Figure 7). However, expression of *hpsA* was not activated in the Δ*sigJ* mutant. Expression of *hpsA* was largely unaffected in the Δ*sigC* strain and showed a moderate reduction in the Δ*sigF* mutant. Expression of *hpsE* was more immediate but also more moderate, peaking at 18 hours in both the wild type and the Δ*sigC* deletion strain. Activation of expression of *hpsE* was lowered in the Δ*sigJ* mutant, but still enhanced in the later developmental stages, indicating that expression of *hpsE* may not be completely knocked out in

the absence of SigJ. Activation of expression of *hpsE* was also lower in the *ΔsigF* mutant, but the difference was not substantial (Figure 7).

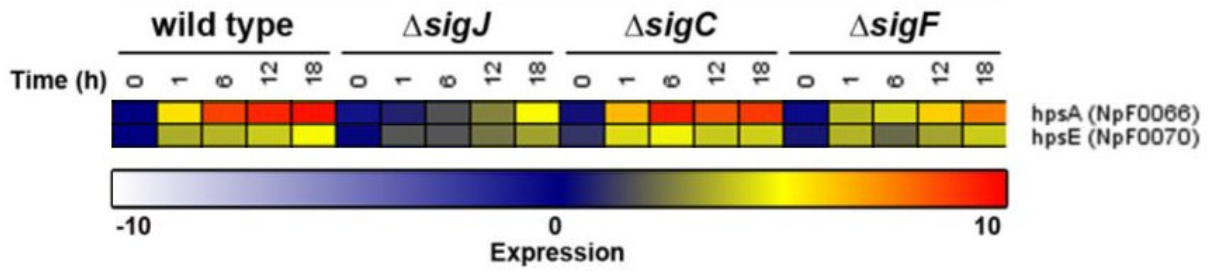

Figure 7: Expression of *hps* genes shown as log2 (average normalised expression for each strain and time point/average normalised expression of the wild-type at 0 h). Heat map showing changes in expression levels of *hpsA* and *hpsE* versus the wild type strain at 0 hours measured using log_2 transformed values.

Cell Division in Hormogonia is Modulated by Expression of *sigC* While Cell Morphology is Dependent on Expression of *sigJ*

Upon induction of vegetative cells into hormogonia, reductive cell division occurs, and the ends of each filament become tapered. Expression of cell division genes was measured by analysis of expression of *ftsQ* and *ftsZ*, which are known to be involved in cell division, and show a transient increase in expression in the wild-type strain between 1-6 hours post-induction. Expression of these cell division genes was affected in both Δ*sigJ* and Δ*sigC* deletion strains, although the impact of deletion of *sigJ* was less than that of deletion of *sigC*, in which expression was abolished (Figure 8A). These results are consistent with cell size measurements, in which average cell sizes for Δ*sigC* and Δ*sigJ* hormogonia are larger than those of the wild type at 24 hours (Figure 8B). Changes in cell morphology, involving the homolog of the *Escherichia coli* cell shape determining protein MreC [42], also occur during hormogonium development, and expression of *mreC* is activated immediately following induction in the wild type and peaking at 12 hours (Figure 8A). While expression in the Δ*sigC* or the Δ*sigF* mutants showed little change

in expression levels of *mreC*, expression was abolished in the Δ*sigJ* mutant, indicating that *mreC* is under transcriptional control by SigJ.

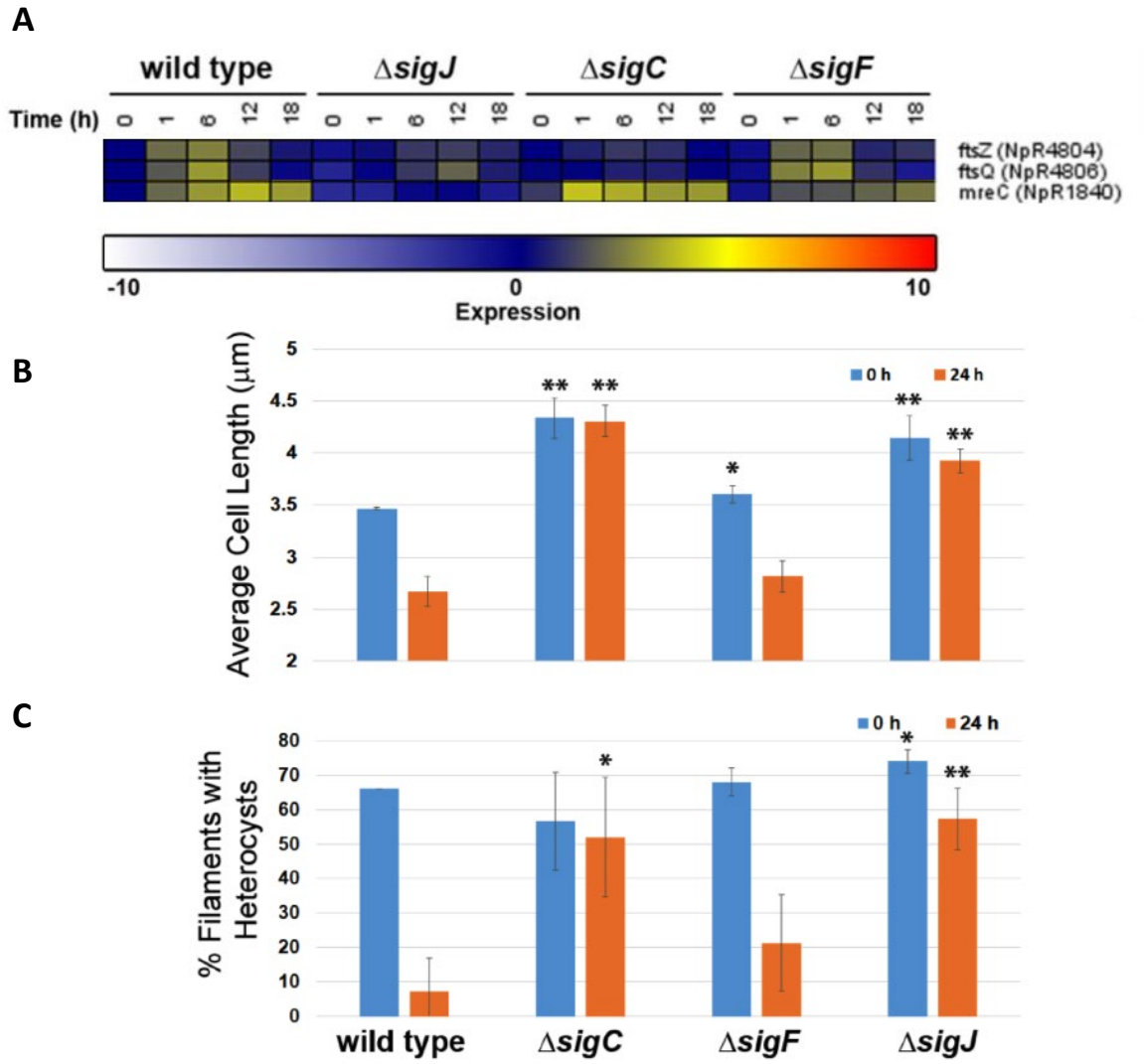

Figure 8: Effects of sigma factor deletions on cell division and cytoskeleton genes. (A) RT-qPCR data showing transcription levels of *ftsZ* (NpR4804), *ftsQ* (NpR4806), and *mreC* (NpR1840) in *N. punctiforme* and its mutants measuring differences in expression levels versus the wild type at 0 hours using $\log_2$ transformed values. Cell division genes were not strongly activated in either the Δ*sigC* or the Δ*sigJ* mutant. (B) Average cell lengths of cells before (blue) and 24 hours after (orange) induction of hormogonium differentiation. (C) Per cent filaments containing heterocysts before and after induction of hormogonium differentiation. * = p-value<0.05, ** = p-value<0.01 as determined by a two-tailed Student's t-Test between the wild-type and each deletion strain at the corresponding time point.

Another aspect of hormogonium development is the fragmentation of the filaments at junctions between vegetative cells and heterocysts, leading to the loss of heterocysts in the resulting hormogonia[4]. This was observed in wild-type *N. punctiforme*, in which a great majority of vegetative filaments contained heterocysts while the presence of filaments in wild-type hormogonia was less than 10 percent (Figure 8C). A similar effect was observed in the Δ*sigF* mutant, although the presence of hormogonium filaments containing heterocysts was approximately 20 percent. In the Δ*sigC* and Δ*sigJ* mutants, the percentage of filamentous containing heterocysts was similar before and after hormogonium induction, indicating that these mutants do not dismember their heterocysts as typically occurs in wild-type hormogonia.

RNAseq Analysis

RNAseq analysis was conducted in conjunction with RT-qPCR to confirm the RT-qPCR data (Figure 9A) and to define the expanded regulons for each sigma factor (Figure 9B). Based on the global transcriptomic data for all upregulated genes during hormogonium development, the regulon for SigJ appears to be the largest of the three, directing expression of over half of the upregulated genes. SigC has a smaller regulon but still directs expression of almost a fourth of the upregulated genes during hormogonium development. SigF has a regulon that appears to be very limited in which it is known to direct expression of *pilA* and possibly a handful of other genes that are upregulated during hormogonium development. One-sixth of the genes that were upregulated during hormogonium development did not involve any of the three sigma factors.

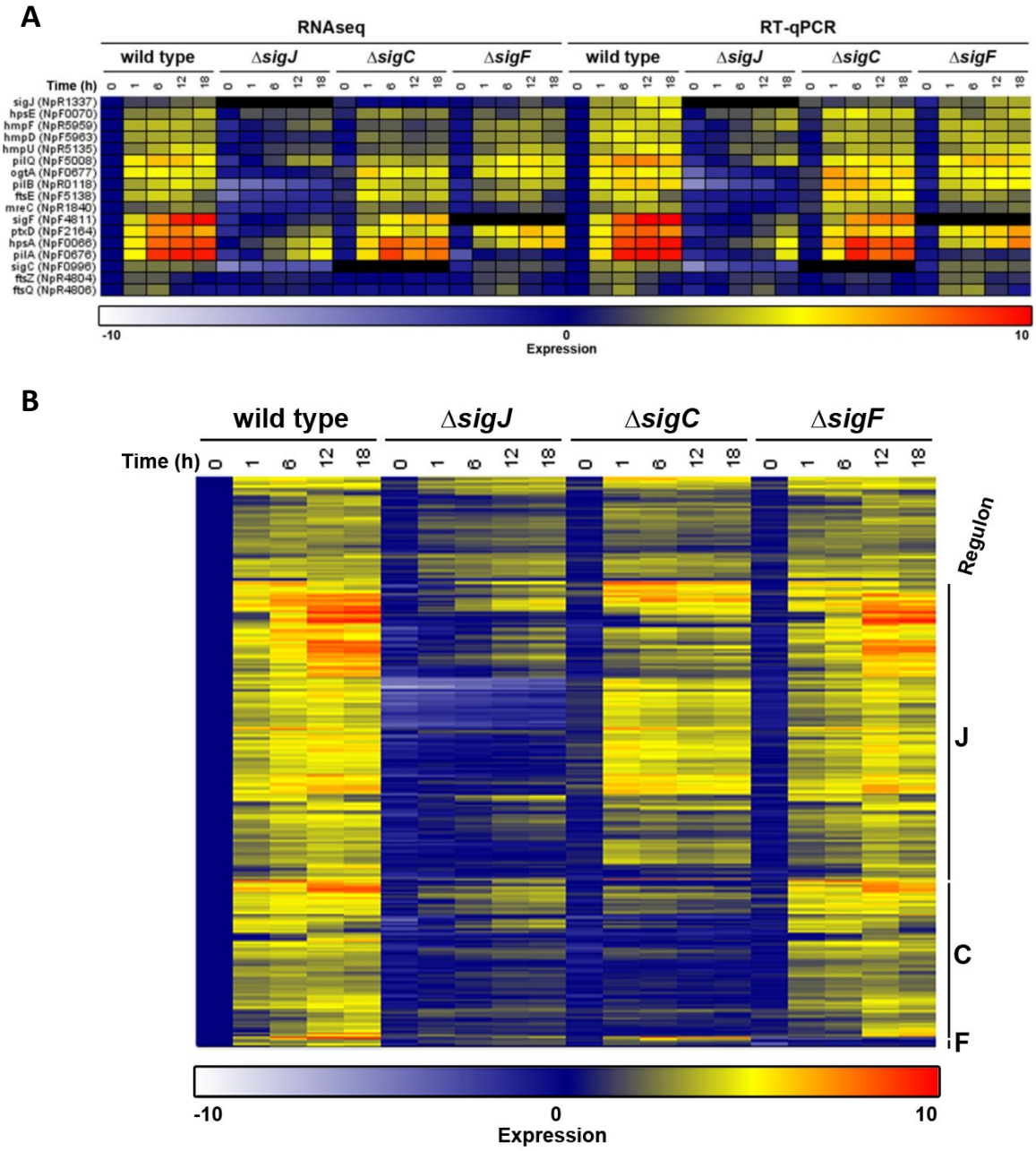

Figure 9: Gene transcription as measured by both RT-qPCR and RNAseq. Expression = log2(average normalised expression for each strain and time point/average normalised expression of the wild-type at 0 h). A) A comparison of RT-qPCR and RNAseq data. (B) Clustered global transcriptome of all upregulated genes during hormogonium development along with the sizes of each of the regulons for each sigma factor.

Chapter 4: Discussion

The results from the application of RT-qPCR and RNAseq supports a model of a hierarchal sigma factor cascade which is activated in hormogonium development. This cascade involves SigJ, as the primary sigma factor and the other sigma factors, *sigC* and *sigF*. SigJ not only activates *sigC* and *sigF*, but it also activates a substantial array of genes including those involved in the type IV pilus system and the chemotaxis-like gene clusters, *hmp* and *ptx*, which are needed to activate secretion of HPS and phototaxis respectively. SigJ also drives expression of genes involved in cell architecture, allowing for the tapered cells at each filament terminus. This data suggests that the regulon of SigJ is large and encompasses a wide variety of different genes, implicating it as a potential master regulator and therefore may be the primary sigma factor of hormogonium development.

RT-qPCR data suggests that SigC is needed to drive expression primarily of cell division genes, which is further supported by cell size measurements of the *sigC* deletion strains versus the wild type strain. Additionally, because enhanced expression of *sigJ* in the sigC-deletion mutant was lost, this suggests that *sigC* may have a role in promoting further expression of *sigJ* and may indicate the possible presence of a positive feedback loop on *sigJ* by SigC activity. These findings suggest a *sigC* regulon which includes cell division genes. Interestingly, many of the SigJ-dependent genes saw a moderate reduction in transcription in the Δ*sigC* mutant, which was observed both using RT-qPCR and RNA-seq. These findings further support a role for SigC in modulating expression and activity of *sigJ* via a positive feedback loop. SigC may potentially direct expression of regulatory proteins that control the activity of the *sigJ* gene product.

Another possibility is that SigC may direct transcription of small interfering RNAs (siRNA) that may inactivate expression of proteins that could block the transcription and activity of *sigJ.* Further investigation needs to be done to fully describe the targets of SigC and the regulatory mechanism involving SigC on *sigJ*.

Compared to SigJ and SigC, the regulon of SigF appears to be very limited. Using RT-qPCR, SigF was found to only direct expression of *pilA*, which encodes the subunits that are polymerised into a pilus by the motor protein complex of the T4P secretion system. More importantly, RNAseq analysis shows that *pilA* was one of the few genes found to be reduced whenever *sigF* expression was absent, supporting the conclusion that *sigF*'s regulon is limited. Even so, like in Δ*sigC*, a few genes that appeared to be dependent on SigJ, such as *hpsA* and *ptxD,* were somewhat reduced, but not completely abolished in Δ*sigF*, potentially indicating that SigF might play a regulatory role on other genes outside its regulon like with SigC. Further investigation on the regulon of SigF needs to be done as well to fully understand the relationships between these three sigma factors during hormogonium development.

Some genes that were activated upon hormogonium development appear to be outside the control of the three sigma factors, as expression of such genes were not completely abolished, even in the Δ*sigJ* mutant. The *hps* locus in particular, although impacted by the deletion of *sigJ*, was not completely abolished, and its expression recovered to around wild type levels in late hormogonium development. This suggests that SigJ is important for expression of such genes, but the influence it exerts is not direct. As explained earlier, this is also the case with the role of SigC and SigF on SigJ-directed genes. This leads to the possibility of genes encoding regulatory products upstream of *sigJ* and the gene products exerting some form of control on expression of these downstream targets. Finally, one of these activated genes may potentially be involved in

activating *sigJ*, causing the initiation of the sigma factor cascade, and all downstream targets leading to the development of motility and other hormogonium characteristics.

We propose a gene regulatory network involving a hierarchal sigma factor cascade, in which *sigJ* is first activated, followed by activation of *sigC* and *sigF*, and other downstream targets by *sigJ* (Figure 10). However, because not all genes that are normally upregulated in hormogonia were completely abolished in any of the sigma factor deletion strains and because we still do not know what activates *sigJ*, there are likely still uncharacterised genes upstream of it that are important in hormogonium development. Therefore, to further understand hormogonium development in filamentous cyanobacteria and what triggers their development, the activation of *sigJ* and its mechanism needs to be investigated, along with the characterisation of activated hormogonium genes outside the sigma factor cascade.

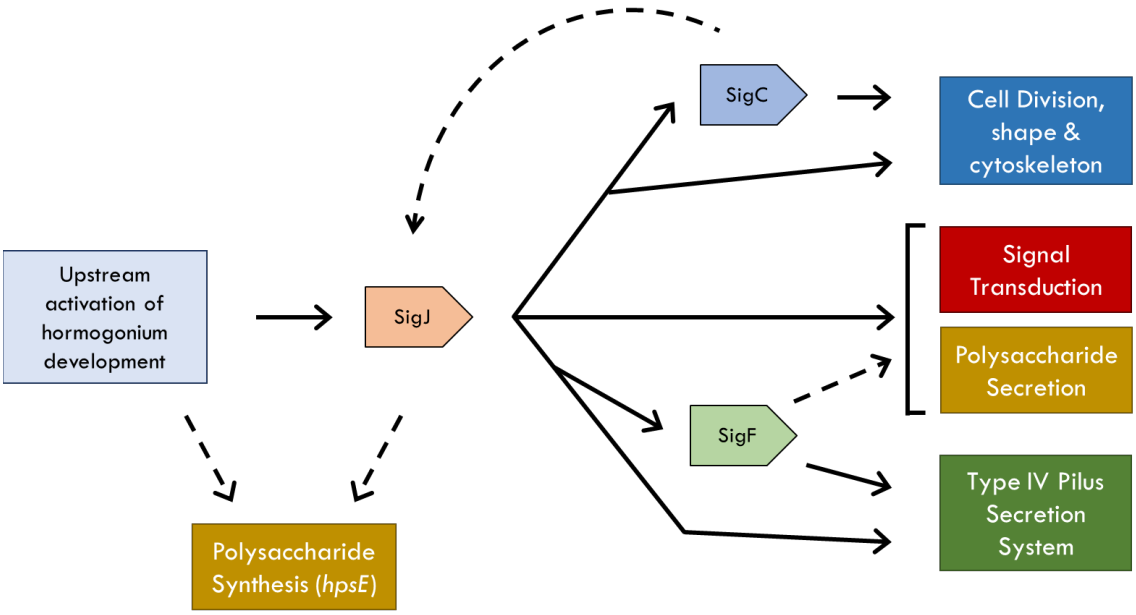

Figure 10: A proposed model of the gene regulatory network controlling hormogonium development. SigJ activates most downstream targets directly, or via activation of downstream sigma factors, and activation of these targets are all associated with hormogonium development. Deletion of *sigJ* led to reduced levels of *hpsE*, but was not abolished, indicating that activation of *hpsE* is done so by an unknown upstream protein. Solid lines indicate direct control while dashed lines indicate indirect control.

These findings are in contrast with previous studies of sigma factors and their involvement in development. Previous studies have investigated development of filamentous cyanobacteria, with many studies using *Anabaena/Nos*toc 7120 as the model organism for development. Because this particular strain was not fully capable of differentiating into all possible cell types of filamentous cyanobacteria, particularly hormogonia, and in fact may have lost it through unintended mutations from prolonged culturing, it was impossible to study the full extent of the effects of the sigma factors of filamentous cyanobacteria. By using *N. punctiforme* as a model organism that is fully capable of differentiating into all cell types, we are now capable of describing more comprehensive roles of sigma factors in global gene expression in filamentous cyanobacteria during differentiation. By having RNAseq data (GSE124969) that coincides with the RT-qPCR data, we now have a comprehensive transcriptomic profile of *N. punctiforme* over the course of hormogonium development and the effects of sigma factor deletions on their transcriptomes, which can be further investigated, along with the regulons of each sigma factor. The regulon of SigJ is of particular interest because it is large and *sigJ* is activated early during hormogonium development. Finally, it is also worth investigating the genes that are outside of the sigma factor cascade which were activated during hormogonium development as they may play a role in the activation of *sigJ* and the sigma factor cascade, along with regulatory control on gene expression of other hormogonium-specific genes.

Filamentous cyanobacteria are developmentally complex microorganisms with relatively simple nutritive requirements. Because of these characteristics, filamentous cyanobacteria are a good platform for studying development. By further describing the mechanisms of gene regulation of developmental genes and proposing pathways, the knowledge of gene regulatory networks in filamentous cyanobacteria may have significant applications in biotechnology,

including the engineering of photosynthetic microbes capable of biofuel production and establishing symbiosis with crop plants. Biofuel production using cyanobacteria represents a potential source of a renewable fuel which can limit the need for non-renewable fossil fuels. This also concerns the increasing need to address climate change, as increasing levels of carbon dioxide in our atmosphere may be recycled into biofuels, while also limiting the addition of new greenhouse gases into our atmosphere. Other cyanobacteria have already been engineered to produce biofuels[43], but *Nostoc* hormogonia remain a potential attractive platform for biofuel production due to its non-dividing nature, allowing *Nostoc* to direct most of its energy into biofuel production rather than the accumulation of biomass. Filamentous cyanobacteria may also have practical uses in agriculture. Since filamentous cyanobacteria naturally establish nitrogen-fixing symbiosis with numerous plants and fungi, filamentous cyanobacteria such as *Nostoc* may, in theory, may be genetically engineered to establish symbiosis with important crop plants. As hormogonia are required to establish new symbioses with plant partners, a fully detailed understanding of the hormogonium gene regulatory network is needed in order to pursue such a feat. The ability of filamentous cyanobacteria to establish nitrogen-fixing symbioses can be a good source of fixed nitrogen for crop plants, limiting the need for nitrogen-based chemical fertiliser.

References

1. Doudna, J.A.; Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science (80-.).* **2014**, *346*.

2. Raven, J.A.; Allen, J.F. Genomics and chloroplast evolution: what did cyanobacteria do for plants? The endosymbiont hypothesis is mainstream. *Genome Biol.* **2003**, *4*, 209.

3. Fischer, W.W.; Hemp, J.; Valentine, J.S. How did life survive Earth's great oxygenation? *Curr. Opin. Chem. Biol.* **2016**, *31*, 166–178.

4. Meeks, J.C.; Campbell, E.L.; Summers, M.L.; Wong, F.C. Cellular differentiation in the cyanobacterium Nostoc punctiforme. *Arch. Microbiol.* **2002**, *178*, 395–403.

5. Meeks, J.C.; Elhai, J.; Meeks, J.C.; Elhai, J. Regulation of Cellular Differentiation in Filamentous Cyanobacteria in Free-Living and Plant-Associated Symbiotic Growth States Regulation of Cellular Differentiation in Filamentous Cyanobacteria in Free-Living and Plant-Associated Symbiotic Growth States. *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 94–121.

6. Golden, J.W.; Yoon, H.S. Heterocyst development in Anabaena. *Curr. Opin. Microbiol.* **2003**, *6*, 557–563.

7. Berman-Frank, I.; Lundgren, P.; Falkowski, P. Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Res. Microbiol.* **2003**, *154*, 157–164.

8. Dok, W.; Hart, B.T. Akinete Differentiation in Anabaena Circinalis (Cyanophyta)1. *J. Phycol.* **1996**, *32*, 557–565.

9. Ducat, D.C.; Way, J.C.; Silver, P.A. Engineering cyanobacteria to generate high-value products. *Trends Biotechnol.* **2011**, *29*, 95–103.

10. Johnson, T.J.; Gibbons, J.L.; Gu, L.; Zhou, R.; Gibbons, W.R. Molecular genetic improvements of cyanobacteria to enhance the industrial potential of the microbe: A review. *Biotechnol. Prog.* **2016**, *32*, 1357–1371.

11. Khayatan, B.; Meeks, J.C.; Risser, D.D. Evidence that a modified type IV pilus-like system powers gliding motility and polysaccharide secretion in filamentous cyanobacteria. *Mol. Microbiol.* **2015**, *98*, 1021–1036.

12. Campbell, E.L.; Meeks, J.C. Characteristics of hormogonia formation by symbiotic nostoc spp. in response to the presence of Anthoceros punctatus or its extracellular products. *Appl. Environ. Microbiol.* **1989**, *55*, 125–131.

13. Campbell, E.L.; Hagen, K.D.; Chen, R.; Risser, D.D.; Ferreira, D.P.; Meeks, J.C. Genetic analysis reveals the identity of the photoreceptor for phototaxis in hormogonium filaments of Nostoc punctiforme. *J. Bacteriol.* **2015**, *197*, 782–791.

14. Breathwaite, E.K.; Prasse, C.E.; Elhai, J.; Chiu, W.-L.; Secor, C.R.; Fraley, E.R.; Chibane, F.L.; Khamar, H.J. Multiple Roles of Soluble Sugars in the Establishment of Gunnera-Nostoc Endosymbiosis. *Plant Physiol.* **2010**, *154*, 1381–1389.

15. Splitt, S.D.; Risser, D.D. The non-metabolizable sucrose analog sucralose is a potent inhibitor of hormogonium differentiation in the filamentous cyanobacterium Nostoc punctiforme. *Arch. Microbiol.* **2016**, *198*, 137–147.

16. Shatwell, T.; Köhler, J. Decreased nitrogen loading controls summer cyanobacterial blooms without promoting nitrogen-fixing taxa: Long-term response of a shallow lake. *Limnol. Oceanogr.* **2019**, *64*, S166–S178.

17. Hospenthal, M.K.; Costa, T.R.D.; Waksman, G. A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nat. Rev. Microbiol.* **2017**, *15*, 365–379.

18. Wilde, A.; Mullineaux, C.W. Motility in cyanobacteria: Polysaccharide tracks and Type IV pilus motors. *Mol. Microbiol.* **2015**, *98*, 998–1001.

19. Risser, D.D.; Chew, W.G.; Meeks, J.C. Genetic characterization of the hmp locus, a chemotaxis-like gene cluster that regulates hormogonium development and motility in Nostoc punctiforme. *Mol. Microbiol.* **2014**, *92*, 222–233.

20. Cho, Y.W.; Gonzales, A.; Harwood, T. V.; Huynh, J.; Hwang, Y.; Park, J.S.; Trieu, A.Q.; Italia, P.; Pallipuram, V.K.; Risser, D.D. Dynamic localization of HmpF regulates type IV pilus activity and directional motility in the filamentous cyanobacterium Nostoc punctiforme. *Mol. Microbiol.* **2017**, *106*, 252–265.

21. Wadhams, G.H.; Armitage, J.P. Making sense of it all: Bacterial chemotaxis. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 1024–1037.

22. Riley, K.W.; Gonzalez, A.; Risser, D.D. A partner-switching regulatory system controls hormogonium development in the filamentous cyanobacterium Nostoc punctiforme. *Mol. Microbiol.* **2018**, *109*, 555–569.

23. Wösten, M.M.S.M. Eubacterial sigma-factors. *FEMS Microbiol. Rev.* **1998**, *22*, 127–150.

24. Davis, M.C.; Kesthely, C.A.; Franklin, E.A.; MacLellan, S.R. The Essential Activities of the Bacterial Sigma Factor. *Can. J. Microbiol.* **2017**, *99*, 1–36.

25. Imamura, S.; Asayama, M. Sigma factors for cyanobacterial transcription. *Gene Regul. Syst. Biol.* **2009**, *3*, 65–87.

26. Haugen, S.P.; Berkmen, M.B.; Ross, W.; Gaal, T.; Ward, C.; Gourse, R.L. rRNA Promoter Regulation by Nonoptimal Binding of σ Region 1.2: An Additional Recognition Element for RNA Polymerase. *Cell* **2006**, *125*, 1069–1082.

27. Paget, M.S. Bacterial sigma factors and anti-sigma factors: Structure, function and

distribution. *Biomolecules* **2015**, *5*, 1245–1265.

28. Aldea, M.R.; Mella-Herrera, R.A.; Golden, J.W. Sigma factor genes sigC, sigE, and sigG are upregulated in heterocysts of the cyanobacterium Anabaena sp. strain PCC 7120. *J. Bacteriol.* **2007**, *189*, 8392–8396.

29. Yoshimura, H.; Okamoto, S.; Tsumuraya, Y.; Ohmori, M. Group 3 sigma factor gene, sigJ, a key regulator of desiccation tolerance, regulates the synthesis of extracellular polysaccharide in cyanobacterium Anabaena sp. strain PCC 7120. *DNA Res.* **2007**, *14*, 13–24.

30. Bell, N.; Lee, J.J.; Summers, M.L. Characterization and in vivo regulon determination of an ECF sigma factor and its cognate anti-sigma factor in Nostoc punctiforme. *Mol. Microbiol.* **2017**, *104*, 179–194.

31. Ehira, S.; Miyazaki, S. Regulation of Genes Involved in Heterocyst Differentiation in the Cyanobacterium Anabaena sp. Strain PCC 7120 by a Group 2 Sigma Factor SigC. *Life* **2015**, *5*, 587–603.

32. Bhaya, D.; Watanabe, N.; Ogawa, T.; Grossman, a R. The role of an alternative sigma factor in motility and pilus formation in the cyanobacterium Synechocystis sp. strain PCC6803. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3188–3193.

33. Khayatan, B.; Bains, D.K.; Cheng, M.H.; Cho, Y.W.; Huynh, J.; Kim, R.; Omoruyi, O.H.; Pantoja, A.P.; Park, J.S.; Peng, J.K.; et al. A Putative O-Linked ɴʟ-N- Acetylglucosamine Transferase Is Essential for Hormogonium Development and Motility in the Filamentous Cyanobacterium Nostoc punctiforme. **2017**, *199*, 1–14.

34. Nakao, M.; Okamoto, S.; Kohara, M.; Fujishiro, T.; Fujisawa, T.; Sato, S.; Tabata, S.; Kaneko, T.; Nakamura, Y. CyanoBase: The cyanobacteria genome database update 2010.

Nucleic Acids Res. **2009**, *38*, 2009–2011.

35. Livak, K.J.; Schmittgen, T.D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **2001**, *25*, 402–408.
36. Sturn, A.; Quackenbush, J.; Trajanoski, Z. Genesis: Cluster analysis of microarray data. *Bioinformatics* **2002**, *18*, 207–208.
37. Tjaden, B. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol.* **2015**, *16*, 1–10.
38. McClure, R.; Balasubramanian, D.; Sun, Y.; Bobrovskyy, M.; Sumby, P.; Genco, C.A.; Vanderpool, C.K.; Tjaden, B. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res.* **2013**, *41*, 1–16.
39. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192.
40. Angelini, C.; Cutillo, L.; De Canditiis, D.; Mutarelli, M.; Pensky, M. BATS: A Bayesian user-friendly software for Analyzing Time Series microarray experiments. *BMC Bioinformatics* **2008**, *9*, 1–13.
41. Risser, D.D.; Meeks, J.C. Comparative transcriptomics with a motility-deficient mutant leads to identification of a novel polysaccharide secretion system in Nostoc punctiforme. *Mol. Microbiol.* **2013**, *87*, 884–893.
42. Singh, S.P.; Montgomery, B.L. Determining cell shape: Adaptive regulation of cyanobacterial cellular differentiation and morphology. *Trends Microbiol.* **2011**, *19*, 278–285.
43. Luan, G.; Lu, X. Tailoring cyanobacterial cell factory for improved industrial properties.

Biotechnol. Adv. **2018**, *36*, 430–442.